Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2019) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

Postproceedings of the 10th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2019 (Tenth Annual Meeting of the BICA Society)

Natural logic is diagrammatic reasoning about mental models

John F. Sowa

Abstract

Diagrammatic reasoning, based on visualization and analogy, is the foundation for reasoning in ordinary language and the most esoteric theories of mathematics. Long before they write a formal proof, mathematicians develop their ideas with diagrams, visualize novel patterns, and discover creative analogies. For over two millennia, Euclid's diagrammatic methods set the standard for mathematical rigor. But the abstract algebra of the 19th century led many mathematicians to claim that all formal reasoning must be algebraic. Yet C. S. Peirce and George Polya recognized that Euclid's diagrammatic reasoning is a better match to human thought patterns than algebraic rules and notations. Peirce's graph logic, combined with Polya's heuristics and Euclid's diagrams, is a better candidate for a natural logic than any algebraic formalism. Psychologically, it supports Peirce's claim that his existential graphs (EGs) provide "a moving picture of the action of the mind in thought." Logically, EGs have a formal mapping to and from the ISO standard for Common Logic. Computationally, algorithms for Cognitive Memory (CM) and virtual reality (VR) can support cross-modal analogies between language-like and image-like representations, static or dynamic.

Peer-review under responsibility of the scientific committee of the 9th Annual International Conference on Biologically Inspired Cognitive Architectures. This copy contains revisions and extensions to the published version, including a new Section 5.

Keywords: natural logic; diagrammatic reasoning; mental models; existential graphs; rules of inference; semiotic; artificial intelligence.

1. Diagrammatic reasoning

Charles Sanders Peirce was a pioneer in logic. Although Frege published the first complete system of first-order logic [7], no one else adopted his notation. A few years later, Peirce published the algebraic version of first-order and higher-order logic [18]. With a change of symbols by Peano and some extensions by Whitehead and Russell, Peirce-Peano algebra is still the most widely used logic today [27]. But Peirce also developed graph notations to express "the atoms and molecules of logic." In 1897, he developed *existential graphs* (EGs) as a notation for the semantics of first-order predicate calculus with equality [19]. For propositional EGs (Alpha) and first-order EGs (Beta), Peirce used the same structure and rules of inference in every version from 1897 to 1911. For Gamma graphs, which represent metalanguage, higher-order logic, and modal logic, he added more syntactic and semantic features [25, 28]. For artificial intelligence, EGs can serve as a foundation for a cognitive architecture.

Peirce's writings on logic, semiotic, and diagrammatic reasoning, which had been neglected during the 20th century, are now at the forefront of research in the 21st [29, 34, 35]. Frege's rejection of psychologism and "mental pictures" reinforced the behaviorism of the early 20th century. But the latest work in neuroscience uses "folk psychology" and introspection to interpret data from brain scans [4]. The neuroscientist Damasio [2] explicitly said that brains create "images, the main currency of our minds. Ultimately consciousness allows us to experience maps as images, to manipulate those images, and to apply reasoning to them." The psychologist Johnson-Laird, who had written extensively about mental models, said that Peirce's existential graphs and rules of inference are a good candidate for a neural theory of reasoning" [12]:

> Peirce's existential graphs are remarkable. They establish the feasibility of a diagrammatic system of reasoning equivalent to the first-order predicate calculus. They anticipate the theory of mental models in many respects, including their iconic and symbolic components, their eschewal of variables, and their fundamental operations of insertion and deletion. Much is known about the psychology of reasoning... But we still lack a comprehensive account of how individuals represent multiply-quantified assertions, and so the graphs may provide a guide to the future development of psychological theories.

Diagrammatic reasoning is one of Peirce's most brilliant insights. His observations, quoted below, are consistent with remarks by creative mathematicians [9,10, 26], but not with the assumptions of Frege and Russell:

> All necessary reasoning without exception is diagrammatic. That is, we construct an icon of our hypothetical state of things and proceed to observe it. This observation leads us to suspect that something is true, which we may or may not be able to formulate with precision, and we proceed to inquire whether it is true or not. For this purpose it is necessary to form a plan of investigation, and this is the most difficult part of the whole operation. We not only have to select the features of the diagram which it will be pertinent to pay attention to, but it is also of great importance to return again and again to certain features. [24, 2:212]

> The word *diagram* is here used in the peculiar sense of a concrete, but possibly changing, mental image of such a thing as it represents. A drawing or model may be employed to aid the imagination; but the essential thing to be performed is the act of imagining. Mathematical diagrams are of two kinds; 1st, the geometrical, which are composed of lines (for even the image of a body having a curved surface without edges, what is mainly seen by the mind's eye as it is turned about, is its generating lines, such as its varying outline); and 2nd, the algebraical, which are arrays of letters and other characters whose interrelations are represented partly by their arrangement and partly by repetitions. If these change, it is by instantaneous metamorphosis. [23, 4:219]

> We form in the imagination some sort of diagrammatic, that is, iconic, representation of the facts, as skeletonized as possible. The impression of the present writer is that with ordinary persons this is always a visual image, or mixed visual and muscular... This diagram, which has been constructed to represent intuitively or semi-intuitively the same relations which are abstractly expressed in the premisses, is then observed, and a hypothesis suggests itself that there is a certain relation between some of its parts — or perhaps this hypothesis had already been suggested. In order to test this, various experiments are made upon the diagram, which is changed in various ways. [22, 2.778]

These quotations suggest two new rules of inference: *observation* and *imagination.* They are generalizations of Peirce's EG rules to accommodate arbitrary images. Observation allows any EG that is derived from an image to be inserted or erased in the same context as that image. Imagination allows any image that is described by an EG to be inserted or erased in the same context as that EG. These two rules reverse Frege's notion of psychologism. Logic is not based on psychology. Instead, psychology is based on logic, which is based on mathematics, which is based on diagrammatic reasoning. Peirce stated that point explicitly:

> Diagrammatic reasoning is the only really fertile reasoning. If logicians would only embrace this method, we should no longer see attempts to base their science on the fragile foundations of metaphysics or a psychology not based on logical theory. [22, 4.571]

Mental images are memories, reconstructions, or transformations of images from the senses. For board games like chess, diagrammatic reasoning is the essence of the game. Most chess experts can play a good blindfold game. For them, the board and pieces are the equivalent of Peirce's "drawing or model," which is a helpful, but optional aid to the imagination. To study the thought processes of mathematicians, Hadamard [9] asked some of the most creative to answer a few questions. Their responses support the observations by Peirce and Polya. Einstein even used Peirce's words *visual* and *muscular*:

> The words or the language, as they are written or spoken, do not seem to play any role in my mechanism of thought. The psychical entities which seem to serve as elements in thought are certain signs and more or less clear images which can be voluntarily reproduced and combined... The above-mentioned elements are, in my case, of visual and some of muscular type. Conventional words or other signs have to be sought for laboriously only in a secondary stage, when the mentioned associative play is sufficiently established and can be reproduced at will.

There is a continuum from counting sticks and drawings in the sand to the most elaborate mathematics. That continuity, according to Peirce, implies that the diagrammatic reasoning of mathematics is continuous with formal and informal reasoning in every field of science, engineering, business, and everyday life. He used that principle to classify the sciences [20, 21]. The dotted lines in Figure 1 show how sciences on the lower right depend on theories developed by sciences on the upper left. All theories of any kind are derived by diagrammatic reasoning.

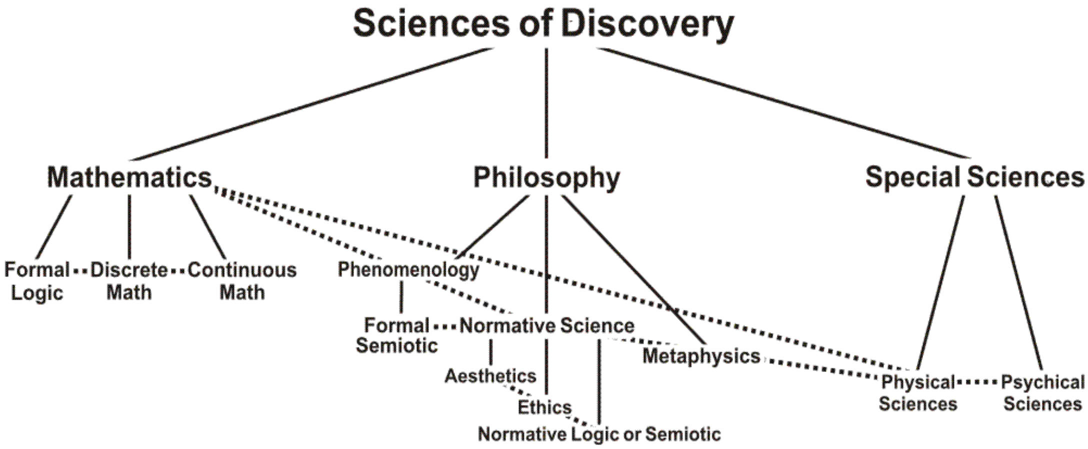


Figure 1: Peirce's classification of the sciences

Logic appears in two places in Figure 1. Formal logic is a branch of pure mathematics. Normative logic (how people should reason) is based on mathematics, phenomenology, aesthetics, and ethics. The special sciences depend on metaphysics, but they also adopt and apply any theory of mathematics that may be useful. Experience consists of mental images and feelings of any kind. Phenomenology is an application of diagrammatic reasoning to analyze and classify experience in signs and relations among signs. The result is *formal semiotic*. Normative science is an application of formal logic and semiotic to evaluate judgments about Beauty, Goodness, and Truth. Peirce considered normative logic and normative semiotic as two aspects of the same subject.

Peirce's classification raises issues that have been debated for centuries. Plato claimed that mathematical forms (diagrams) exist prior to any physical embodiment, but Aristotle maintained that the forms are abstractions from physical things. Aristotle's claim is based on his semiotic: "First we must determine what noun (*onoma*) and verb (*rhêma*) are, and after that what negation, assertion, proposition, and sentence. These spoken sounds are symbols of mental experiences (*pathêmata tês psychês*), and the written letters are symbols of the sounds" [1]. The medieval Scholastics developed semiotic in detail [3], Peirce refined their theories with modern logic, and C. I. Lewis, who was strongly influenced by Peirce, introduced the term *qualia* for those mental experiences [15].

To resolve the debates, Peirce postulated three universes of discourse: the possible, the actual, and the necessary. Mathematics is the universe of all possible diagrams and theories about them. The special sciences use those theories to analyze the actual world. Mathematics applied to actuality makes testable predictions. Hypotheses that make reliable predictions become laws of science. The totality of laws about anything actual or possible is the universe of the necessitated. The dotted lines of Figure 1 transmit signs in both directions. People discover new diagrams as abstractions from images of the actual. Those diagrams become hypotheses for mathematical theories. Then the special sciences use those theories to make further observations of the actual. Lewis's word *qualia* is useful for relating Peirce's semiotic to modern cognitive science.

2. Existential Graphs

For Peirce, algebra is a useful linearization for documenting mathematical patterns. But diagrams in two or more dimensions can show the patterns more directly. They can provide deeper insights into the mathematical structures and their mapping to and from structures in the world. For existential graphs, Peirce discovered elegant, notation-independent rules of inference that can be applied to diagrams in any number of dimensions. Those rules are a refinement and generalization of *natural deduction* by Gentzen [8], but Peirce discovered them thirty years earlier. With minor adjustments for syntax, the rules can be applied to any notation that can be translated to first-order or higher-order logic. Peirce even thought of adapting them to "stereoscopic moving pictures" [23, 3:191]. That was his term for today's virtual reality.

In 1911, Peirce's wrote his clearest, simplest, and most general specification for EGs [23, 3:162-169]. All quoted sentences in this section are from those pages. As an example, the EG on the left of Figure 2 asserts that there is a phoenix. The line, called a *line of identity*, represents existence. It asserts *There is something*. The word *phoenix* is the *name* of a monadic relation. A line attached to that name asserts *There is something of type phoenix*. For the EG in the middle, Peirce explained "To deny that there is any phoenix, we shade that assertion which we deny as a whole. Thus what I have just scribed means 'It is false that there is a phoenix.' But the [EG on the right] only means 'There is something that is not identical with any phoenix'."

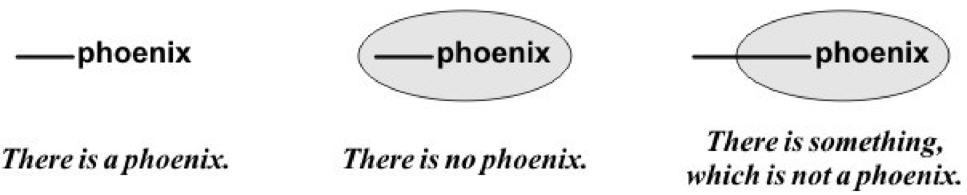


Figure 2. Three existential graphs

In predicate calculus notation, the EGs in Figure 2 would be written

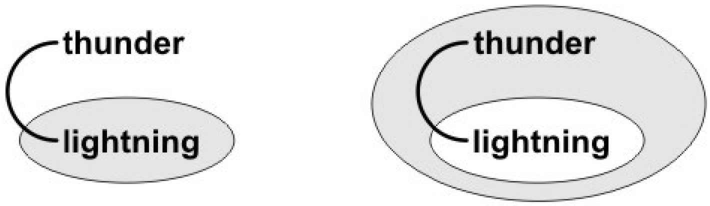
```
PC Left:     ∃x  phoenix(x)
PC Middle:   ~∃x  phoenix(x)
PC Right:    ∃x  ~phoenix(x)
```

In the Common Logic Interface to Predicate calculus (CLIP), they would be written

```
Left:        (∃x) (phoenix x).
Middle:      ~[ (∃x) (phoenix x) ].
Right:       (∃x)  ~[ (phoenix x) ].
```

In CLIP, every pair of parentheses or brackets marks a *node*, which may be inserted, copied, or erased by the EG rules of inference. The *quantifier node* **(∃x)** marks the beginning of a line of identity. It asserts the existence of something named **x**. The *relation node* **(phoenix x)** has one *peg*, which is attached to the line **x**. In a relation node, the left parenthesis is placed in front of the relation name. For a *negation node* (oval), square brackets mark the boundary between a shaded area (negative) and an unshaded area (positive). The tilde **~** in front of the bracket is a reminder of the negation, not a requirement. In certain contexts, the tilde may be replaced by a keyword, such as **If**, **Then**, **Or**.

When an oval is drawn inside another oval, the doubly nested area is positive (unshaded), as in Figure 3. Any area nested inside an odd number of ovals is shaded, and any area inside an even number of ovals (possibly zero) is unshaded. As Peirce said, “The graph [on the left] asserts that it thunders without lightning... a denial shades the unshaded and unshades the shaded. Consequently [the graph on the right] means ‘If it thunders, it lightens’.”



Something thunders, and it does not lighten. ***If something thunders, then it lightens.***

Figure 3. An EG with one negation and an EG with two negations

In predicate calculus, all Boolean operators may be expressed with just negations and conjunctions. As an option, an if-then operator may be represented with an arrow → or with Peano’s symbol ⊃ . But that option requires the quantifier **∃x** to be moved to the front and be replaced by the universal **∀x**:

```
PC Left:      ∃x  (thunder(x)  ∧  ~lightning(x))
PC Right:     ~∃x  (thunder(x)  ∧  ~lightning(x))
PC Optional:  ∀x  (thunder(x)  → lightning(x))
```

In EGs, conjunction is implicit, and there is no ordering of nodes in any area. In CLIP, the symbol ∧ is omitted, and all permutations of nodes are logically equivalent. But a quantifier node, such as **(∃x)**, must occur before (to the left of) any occurrence of the name **x** in any other node. In CLIP, Option 1 below allows **(∃x)** in the if-area to govern **(lightning x)** in the then-area. To make CLIP look more like predicate calculus, Option 2 moves the quantifier node **(∃x)** to the front and converts it to **(∀x)**:

```
Left:        (∃x)  (thunder x)  ~[ (lightning x) ].
Right:       ~[ (∃x)  (thunder x)  ~[ (lightning x) ].
Option 1:    [If   (∃x)  (thunder x)  [Then   (lightning x) ] ].
Option 2:    (∀x)  [If  (thunder x)  [Then   (lightning x) ] ].
```

Figure 4, Peirce said, “shows a graph instance composed of instances of three indivisible graphs which assert ‘there is a male’, ‘there is something human’, and ‘there is something African’. The syntactic junction or point of *teridentity* asserts the identity of something denoted by all three.”

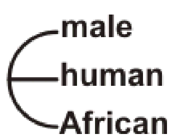


Figure 4. There is a male human African

In CLIP, Figure 4 may be represented by three lines of identity and a *coreference node* **(= x y z)** that indicates a junction of the three lines. A coreference node represents a *ligature* of lines that refer to the same thing. For this example, all occurrences of **y** and **z** may be replaced by **x**. Then the coreference node **(= x x x)** is unnecessary, and it may be deleted. Following are four equivalent ways of translating Figure 4 to CLIP:

Teridentity: (∃x) (male x) (∃y) (human y) (∃x) (African z) (= x y z).
Option 1: (∃ x y z) (male x) (human y) (African z) (= x y z).
Option 2: (∃ x:male y:human z:African) (= x y z).
One name: (∃x) (male x) (human x) (African x).

The first line uses a separate quantifier node for each branch. Option 1 combines all three quantifier nodes in a single node **(∃ x y z)**. Option 2 combines each name in the quantifier with the monadic relation that restricts the range of the quantifier. That option may be read "There is a male x, a human y, and an African z, and all three are the same individual." The last line uses just one name x to say "There is a male human African."

In modern terminology, Peirce's indivisible graphs are called *atoms*. Each atom is represented by its relation name followed by its list of *arguments*, which Peirce called *logical subjects*. In the graphic form, an N-adic relation has N *pegs*, each of which is attached to a line of identity for its logical subject. In CLIP, each atom is represented by a pair of parentheses that enclose a relation name followed by a list of names for its logical subjects. When an EG is translated to CLIP, the labels that correspond to each peg are listed in the order of the pegs, as viewed from left to right. When Peirce drew EGs, he was consistent in preserving that order.

Peirce represented a *proposition* as a relation with zero pegs. He called it a *medad*; the prefix *me-* comes from the Greek *mêden* for *nothing*. In CLIP, a proposition p is represented as a relation with no logical subjects: **(p)**. In early versions of EGs, Peirce distinguished two subsets: Alpha for propositional logic and Beta for first-order logic. By treating medads as relations, he avoided the need to distinguish Alpha from Beta, since the same semantics and rules of inference apply to both.

With these examples, Peirce presented EG syntax in less than three printed pages. The basic EG notation can represent full first-order logic with identity by combining four primitives: (1) existence by a line of identity; (2) assertion by a relation name attached to zero or more lines; (3) negation by a shaded oval; (4) implicit conjunction of all the graphs on a *sheet of assertion* or inside any oval nested in the sheet. The basic CLIP notation represents each of those operators: (1) a quantifier node, such as **(∃x)** to say that something named **x** exists; (2) a relation node with a name of the relation type followed by a list of zero or more names for its l*ogical subjects*; (3) square brackets with a tilde **~[ ]** to represent negation; (4) implicit conjunction of all the nodes in any area.

To improve readability and to simplify the mapping to and from other notations, CLIP includes features that are defined by combinations of the basic four. Figure 5 shows three composite operators defined in terms of nested negations. Even though those operators are composite, the graphical patterns are just as readable as the special symbols used in linear notations..

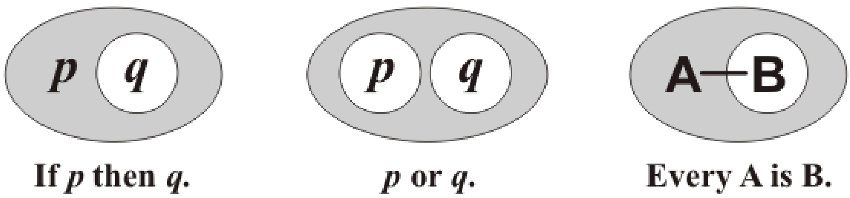


Figure 5. Three composite operators

Another advantage of the graphs is their clarity in showing the scope of quantifiers. In any nest of ovals, the quantifiers (lines of identity) in an outer area govern any nested area. The nested brackets of CLIP also show the scope of quantifiers. But the PC symbols ∨ and → don't show how they affect the scope. That lack of clarity is a frequent cause of errors by students who are studying logic. Following is the CLIP for the left EG of Figure 5:

Basic: ~[(p) ~[(q)]].
Extended: [If (p) [Then (q)]] .

For the middle EG:

```
Basic:      ~[ ~[ (p) ]  ~[ (q) ] ].
Extended:   [Or [ (p) ]  [ (q) ] ].
```

For the right EG, note that the basic CLIP may be translated to several equivalent versions of extended CLIP. The last line may be read "Every x that is A is also B" or just "Every A is a B."

```
Basic:      ~[ (∃x)  (A x)  ~[ (B x) ] ].
Extended:   [If  (∃x)  (A x)  [Then  (B x) ] ].
Extended:   (∀x)  [If  (A x)  [Then  (B x) ] ].
Extended:   [If  (∃ x:A)  [Then  (B x) ] ].
Extended:   (∀ x:A)  [If  [Then  (B x) ] ].
Note the double negation:  (∀ x:A)  ~[  ~[ (B x) ] ].
Erase the double negation: (∀ x:A)  (B x).
```

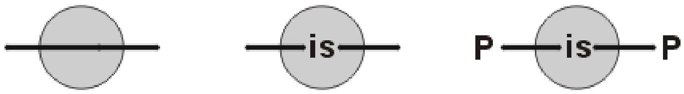


Figure 6. Stating that two things are not identical

In Figure 6, each of the three EGs has two lines of identity that begin outside the shaded area and are extended inside. If the two lines are named x and y, the ligature where they meet would be **(= x y)**. The CLIP for the EG on the left would be **(∃ x y) ~[(= x y)]**. in English, it may be read "There exist an x and a y that are not identical" or "There exist at least two things."

But a straight line that passes through an oval does not shows the point where the ligature is made. One option is draw a heavy dot at the juncture. The EG in the middle shows another option: insert a dyadic relation named is, and define **(is x y)** as a synonym for **(= x y)**. The EG on the right attaches a monadic relation named P at each end. In CLIP, it may be represented **(∃ x:P y:P) ~[(= x y)]**; in English, "There exist two different things with the property P" or simply "There are two Ps."

In summary, CLIP has the same primitives as EGs, but with extra syntax to state the graphic links. An oval for negation is represented by a tilde **~** followed by a pair of square brackets to enclose the CLIP for the nested subgraph. Lines of identity are represented by defining labels and bound labels. The graphic options for connecting and extending lines are shown by coreference nodes. A ligature is represented by a coreference node with two or more names. Some coreference nodes may be eliminated by renaming the lines, as in Figure 4. For more examples, see the tutorial [32]. For the formal details, see the longer article [33].

3. Rules of observation and imagination

For diagrammatic reasoning, Peirce considered images of any kind in any sensory modality [35]. He said that a portrait or drawing, by itself, could represent a possibility. But when a name, such as *George Washington*, is linked to the portrait, the combination asserts the proposition that the possible individual is actually George Washington. The rule of *observation* allows any diagram that expresses a proposition to be replaced by an EG that states the proposition. The rule of *imagination* allows any EG to be replaced by a diagram that is described by the EG. As an example, Figure 7 shows an EG with nested Euclidean diagrams on the left and an EG with nested CLIP on the right. The CLIP names are derived from Euclid's letters and combinations of letters.

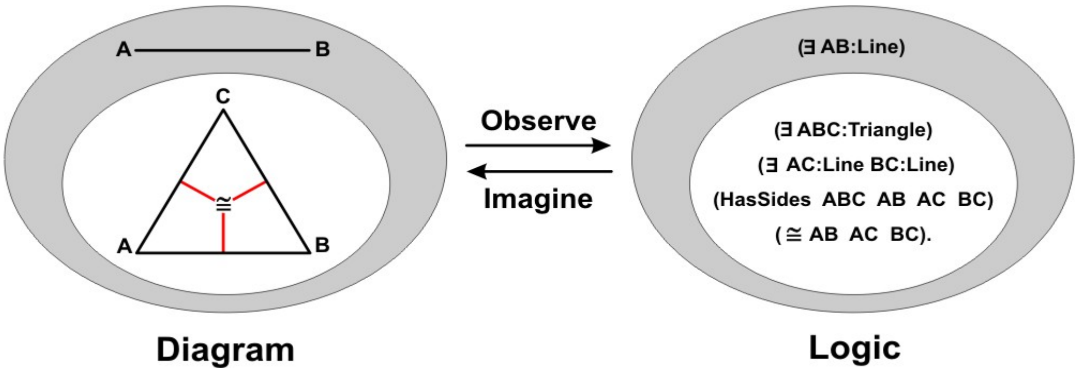


Figure 7: Mapping Euclid's diagrams to and from CLIP

Both EGs in Figure 7 state Proposition 1 in Euclid's *Elements*, which Heath translated "On a given finite straight line, to draw an equilateral triangle" [5]. The CLIP on the right was derived by observing the drawing on the left. The CLIP may be read as the sentence "If there is a line AB, then there is a triangle ABC, which has sides AB, BC, and AC, and the lines AB, AC, and BC are congruent." The shaded area on the left contains a line AB. By iteration, it was copied to the base of the triangle in the white area. By observation, it was mapped to the CLIP sentence **(Line *AB)**, which appears in the shaded area on the right. By imagination, the CLIP sentences on the right may be mapped back to the images on the left. Both EGs represent exactly equivalent logic.

Figure 7 clarifies the role of phenomenology in the classification of Figure 1. The arrows for observation and imagination map mental experience to and from patterns of signs. Formal semiotic analyzes and classifies signs and their combination in diagrams of any kind. Feelings are the basis for value judgments about Beauty, Goodness, and Truth. The addition of value judgments to formal logic and semiotic produces normative logic and semiotic.

Peirce discussed observation and imagination in his writings, but he did not call them rules of inference. For diagrams as precisely drawn as Euclid's, both rules can be stated precisely and added to the other EG rules. The combined set of rules specify a sound and complete proof procedure for Common Logic [11], whose semantics is consistent with Peirce's. To state the rules in full generality, some definitions are required:

- *Diagrams.* A *diagram* is linear, graphic, or blank. A *linear diagram* is a string of one or more symbols. A *graphic diagram* is a combination of at least one symbol and one or more icons and images in any number of dimensions. A *blank diagram* has no symbols, icons, or images.
- *Generalization.* A diagram D is a *generalization* of a diagram S if every case (model) for which S is true, D is also true. In particular, a blank diagram is always true. Therefore, the blank is a generalization of every diagram, including itself.
- *Specialization.* A diagram S is a *specialization* of a diagram D if and only if D is a generalization of S.
- *Equivalence.* Two diagrams D1 and D2 are *equivalent* if each one is a generalization and a specialization of the other. A diagram is equivalent to any copy of itself.

Over the years, Peirce stated his rules of inference in slightly different, but equivalent ways. The version stated below is based on the same text used in Section 2. The rules are grouped in three pairs: one rule (i) inserts a diagram, and the other (e) erases a diagram. The only axiom is the blank, which is always true. The negation of a blank is always false. The rules of inference depend on whether an area is positive (unshaded) or negative (shaded). They do not depend on how deeply nested any area may be. Unlike earlier versions, which were limited to two-dimensional sheets, these rules can be applied to strings, diagrams, images, or models in any number of dimensions. That includes three dimensions plus time for stereoscopic moving images.

- **Insert/erase.** Rule 1i: In a negative area, any diagram D (including the blank) may be replaced by any specialization of D. Rule 1e: In a positive area, any diagram D may be replaced by any generalization of D (including the blank)

- **Iterate/deiterate.** Rule 2i: Any diagram D in any area *c* may be iterated (copied) in the same area *c* or into any area nested in *c*. No diagram may be copied directly into itself. But it is permissible to copy a diagram D in the same area *c* and then copy the copy of D into some area nested inside the original D. Rule 2e: Any diagram D that may have been derived by rule 2i may be erased. Whether or not D had been derived by 2i is irrelevant.
- **Double negation.** Rule 3i: A double negation (nest of two negations with nothing between the inner and outer) may be drawn around any diagram, subdiagram, or set of diagrams in any area. In any area, lines of identity that originate outside the area and pass through it without a connection to anything in the area are not considered part of that area. In CLIP, those lines would not have any bound labels in the area through which they pass. Rule 3e: A double negation in any area may be erased.

For proofs that involve a mixture of notations, either linear or graphic, rule 1i for insertions in a negative context may be allowed to insert diagrams in any notation, symbolic or graphic. Similarly, rule 1e for erasures in a positive context may be allowed to erase diagrams in any notation. The only options that require special treatment are the rules of iteration (2i) and deiteration (2e) in the same area:

- **Observe.** Rule 2i: If a graphic diagram D occurs in any area *c*, any symbolic generalization or equivalence of D may be inserted in *c*.
- **Imagine.** Rule 2i: If a symbolic diagram D occurs in any area *c*, any graphic generalization or equivalence of D may be inserted in *c*.
- **Erase.** Rule 2e: If a diagram D of any kind occurs in any area where any kind of specialization or equivalence of D also occurs, then D may be erased.

If the diagramming conventions are precisely defined, these rules are sound: observation and imagination would add duplicate information in area A; and erasure would delete duplicates. For scenes in nature, photographs, and informal drawings, these rules may be useful, but fallible approximations. For more discussion and examples, see the slides [32] and a more formal article [33].

4. Natural logic

Since the 1970s, psychologists, philosophers, linguists, and logicians have been searching for some innate *natural logic* or *language of thought* [6, 14, 36]. The goal is a theory and notation that could support the full range of human reasoning from the babbling of an infant to the most advanced science. Mental models and diagrammatic reasoning are a good candidate for the semantics, but the syntax of predicate calculus is not a good match to any natural language (NL). Among the many theories of NL syntax and semantics, the Discourse Representation Theory by Kamp and Reyle [13] is widely used in computational linguistics, and their *discourse representation structures* (DRSs) are isomorphic to first-order existential graphs. In fact, a CLIP statement has a more direct mapping to a DRS than to an EG. As an example, Figure 8 shows an EG and DRS for the sentence *If a farmer owns a donkey, then he beats it*. The same CLIP can represent both.

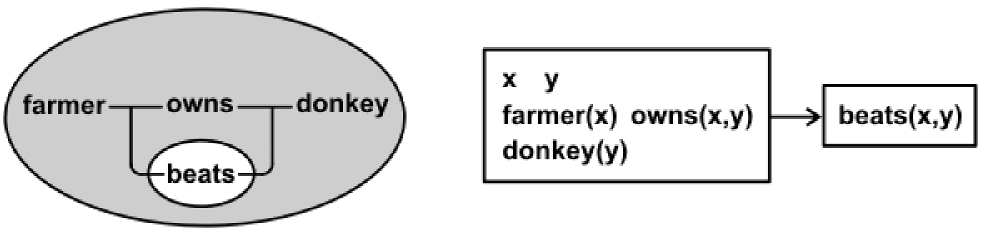


Figure 8. EG (left) and DRS (right) for *If a farmer owns a donkey, then he beats it.*

The DRS and EG primitives are identical: the default *and* operator is omitted, and the default quantifier is existential. DRS negation is represented by a box marked with the ¬ symbol, and implication is represented by two boxes connected by an arrow. As Figure 8 illustrates, nested EG ovals allow lines in the *if* oval to extend into the *then* oval. For DRS, Kamp made an equivalent assumption: the quantifiers for **x** and **y** in the *if* box govern **x** and **y** in the *then* box. Since they are isomorphic, the EG and DRS in Figure 8 can be translated to the same CLIP:

[If (∃ x y) (farmer x) (donkey y) (owns x y) [Then (beats x y)]].

As the translation shows, the unlabeled EG lines are assigned labels **x** and **y** in both DRS and CLIP. The nested negations in the EG may be represented by the keywords **If** and **Then** in the CLIP, and those same keywords may be used to represent the DRS boxes. As another example of the similarity, consider the pair of sentences *Pedro is a farmer. He owns a donkey.* Kamp and Reyle [13] observed that proper names like *Pedro* are not rigid identifiers. In DRS, names are represented by predicates rather than constants. That convention is similar to Peirce's practice with EGs. On the left of Figure 9 are the EGs for the two sentences; on the right are the DRSes.

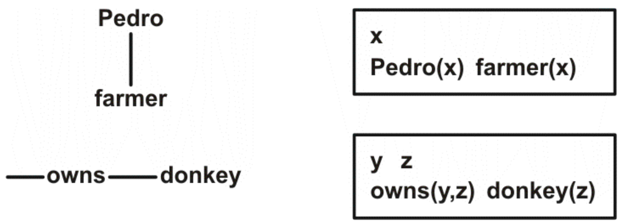


Figure 9. EGs and DRSes for *Pedro is a farmer. He owns a donkey.*

The EGs and DRSs in Figure 9 map to the same two statements in CLIP:

(∃ x) (Pedro x) (farmer x). (∃ y z) (owns y z) (donkey z).

In English, a pronoun such as *he* can refer to something in a previous sentence, but the variable **y** in the second formula cannot refer to the variable **x** in the first. But the EG and DRS notations allow a direct connection. To combine the two EGs, connect the line of identity for *Pedro* to the line for *he* in Figure 9. To combine the two DRSs, transfer the contents of one DRS box to the other, move the list of variables to the top, and insert the equality **x=y**. These operations produce the logically equivalent EG and DRS in Figure 10:

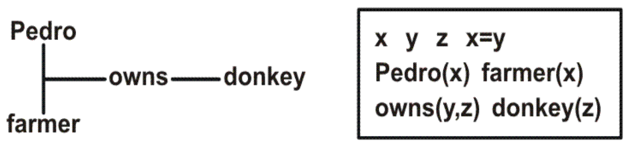


Figure 10. Combining the EGs and DRSs in Figure 9

Since the EG and DRS notations are isomorphic and they're represented by the same CLIP, Peirce's rules of inference can be applied to them in exactly equivalent steps. The proof in Figure 11 begins with the EGs for the sentences *Pedro is a farmer who owns a donkey* and *If a farmer owns a donkey, then he beats it*. Then the rules 2i, 1i, 2e, and 3e derive an EG for the conclusion *Pedro is a farmer who owns and beats a donkey*.

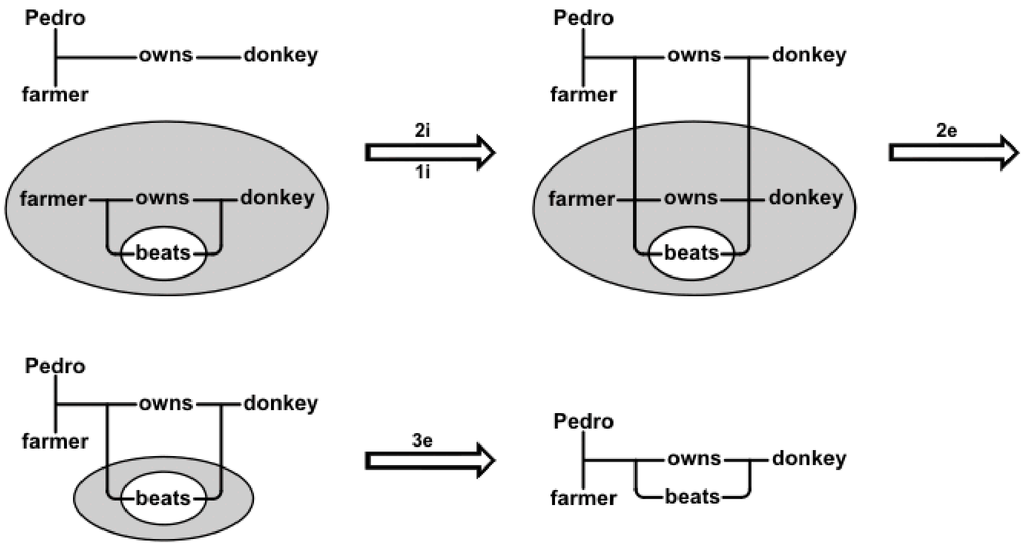


Figure 11. A proof according to Peirce's rules

The proof in Figure 11 takes four steps, but the first arrow combines two steps: extending lines by rule 2i and connecting lines by rule 1i. The following CLIP proof shows all four steps. For the DRS proof, translate each CLIP step to DRS.

1. Starting graphs: To avoid a name clash, the label **x** of the first DRS was replaced by **w**.
   (∃ w z) (Pedro w) (farmer w) (owns w z) (donkey z)
   [If (∃ x y) (farmer x) (donkey y) (owns x y) [Then (beats x y)]].

2. By rule 2i, extend the lines of identity by inserting nodes **(= w)** and **(= z)** into the area of the if-clause:
   (∃ w z) (Pedro w) (farmer w) (owns w z) (donkey z)
   [If (∃ x y) (farmer x) (donkey y) (owns x y) (= w) (= z) [Then (beats x y)]].

3. By 1i, connect the line for **w** to **x** and **z** to **y** with the coreference nodes **(= w x) (= z y)** and relabel:
   (∃ w z) (Pedro w) (farmer w) (owns w z) (donkey z)
   [If (farmer w) (donkey z) (owns w z) [Then (beats w z)]].

4. By 2e, erase the copy of **(farmer w) (donkey z) (owns w z)** in the if-clause:
   (∃ w z) (Pedro w) (farmer w) (owns w z) (donkey z)
   [If [Then (beats w z)]].

5. By 3e, erase the double negation (the if-then pair with an empty if-clause):
   (∃ w z) (Pedro w) (farmer w) (owns w z) (donkey z) (beats w z).

For disjunctions, Figure 12 shows the EG and DRS for the sentence *Either Jones owns a book on semantics, Smith owns a book on logic, or Cooper owns a book on unicorns* [13, p 210]. The EG at the top of Figure 12 shows that the existential quantifiers for Jones, Smith, and Cooper are in the outer area, but the quantifiers for the three books are inside the alternatives.

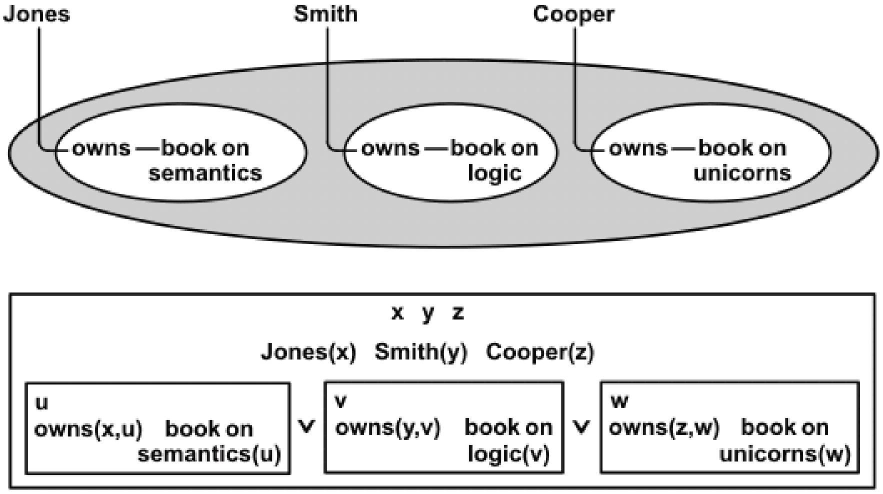


Figure 12. EG and DRS for a disjunction with three alternatives

Both Peirce and Kamp allowed spaces inside relation names, but CLIP requires names with spaces or other special characters to be enclosed in quotes:

```
(∃ x y z) (Jones x) (Smith y) (Cooper z)
~[ ~[ (∃ u) ("book on semantics" u) (owns x u) ]
   ~[ (∃ v) ("book on logic" *v) (owns y v) ]
   ~[ (∃ w) ("book on unicorns" *w) (owns z w) ] ].
```

For better readability, extended CLIP has a keyword **Or** as a synonym for a negation with nested options:

```
(∃ x y z) (Jones x) (Smith y) (Cooper z)
    [Or   ~[ (∃ u) ("book on semantics" u) (owns x u) ]
          ~[ (∃ v) ("book on logic" v) (owns y v) ]
          ~[ (∃ w) ("book on unicorns" w) (owns z w) ] ].
```

The simplicity and generality of EGs, their rules of inference, and their methods for evaluating truth make them a good candidate for a natural logic. To illustrate the generality, the next two examples show that Peirce's rules can be applied directly to English sentences. The only addition is a convention for shading negative areas. In Figure 13, the word *every*, which is on the boundary of an oval, creates the same pattern of nested negations as *if-then*. The word *cat*, by itself, expresses the proposition *There is a cat*. The phrase *cat in the house* expresses the conjunction *There is a cat and it is in the house*. By the rule of specialization (1i), the word *cat* in a negative area may be specialized to the phrase *cat in the house*. By the rule of generalization (1e), the word *carnivore* may be generalized to *animal*.

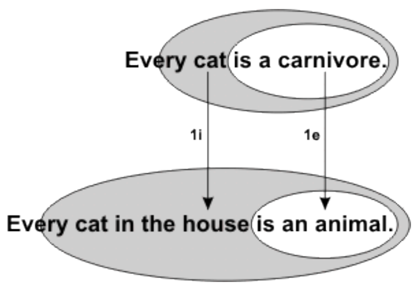


Figure 13. Specializing and generalizing English phrases

Peirce's rules of inference be can directly applied to NL sentences that can be mapped to and from EG or DRS. But some reordering may be needed, especially for highly inflected languages such as Latin or Russian. Figure 14 shows a four-step proof. By specialization (1i), the clause *a cat is on a mat* is specialized to *Yojo is on a mat*. The pronoun *it* represents an EG line of identity or a CLIP coreference node. By iteration (2i), a copy of the name Yojo replaces the pronoun *it*. By deiteration (2e), the clause *Yojo is on a mat* is erased because it is identical to a sentence in the outer area. Finally, the shaded area and the words *if* and *then* on the boundaries of the ovals are erased by the rule of double negation (3e).

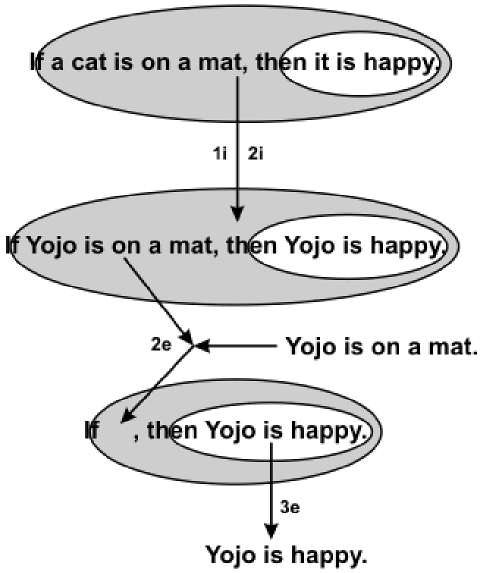


Figure 14. A four-step proof by Peirce's rules

Since Peirce's rules are sound and complete for first-order logic, they are also sound and complete for any subset of NL that can be translated to and from FOL. But Peirce also understood the source of vagueness. Formal precision is a sharp endpoint of a continuum whose vague end disappears in a fog:

> The *vague* might be defined as that to which the principle of contradiction does not apply... But wherever degree or any other possibility of continuous variation subsists, absolute precision is impossible. Much else must be vague, because no man's interpretation of words is based on exactly the same experience as any other man's. Even in our most intellectual conceptions, the more we strive to be precise, the more unattainable precision seems. It should never be forgotten that our own thinking is carried on as a dialogue, and though mostly in a lesser degree, is subject to almost every imperfection of language [22, 5.505-506].

As versions of formal logic, EG and CLIP are as precise. That precision is possible in mapping a formal logic to and from a "skeletonized" diagram such as Euclid's. But Euclidean diagrams cannot represent actual objects with absolute precision. Some degree of vagueness or experimental error is inevitable..

5. Peirce's legacy: A general theory of cognition

Natural logic, mental models, and diagrammatic reasoning are three closely related aspects of cognition. They support the full range of reasoning from informal common sense to the most advanced theories of science. The rules of observation and imagination are the starting point for phenomenology. Observation is the source of evidence about what exists; imagination is the source of ideas about what can or should exist. Imagination is a kind of abduction, which must be tested by deduction, experiment, and analysis of the consequences. As Peirce wrote,

> When a man desires ardently to know the truth, his first effort will be to imagine what that truth can be. He cannot prosecute his pursuit long without finding that imagination unbridled is sure to carry him off the track. Yet nevertheless, it remains true that there is, after all, nothing but imagination that can ever supply him an inkling of the truth. [22, 1.46]

> An imagination is an affection of consciousness which can be directly compared with a percept in some special feature and be pronounced to accord or disaccord with it. [22, 2.148]

For diagrammatic reasoning, Peirce developed existential graphs, which he called "the logic of the future." Although he used observation and imagination to derive the premises for logic, he did not consider those operations rules of inference. But the notation independence of the 1911 EGs enables arbitrary images to be inserted in any area of an EG. In the same text [23, 3:191], Peirce also discussed "thinking in stereoscopic moving images" and the relationship to "real objects moving in solid space." That comment suggests the option of diagrammatic reasoning in any number of dimensions. That would enable continuous, dynamic transformations of mental models.

For formal reasoning, the CLIP linearization is useful for mapping EGs to and from other notations for logic. With some extensions, CLIP has the full expressive power of the ISO/IEC standard for Common Logic [11]. Those extensions include functions, metalanguage, and the option of quantifiers that range over functions and relations. Metalanguage and quantified names for relations are two features of Peirce's Gamma graphs. A function in CLIP is treated as a relation in which the value of the last *peg* or *logical subject* is determined by the values of the logical subjects that precede it. For example, the equation 2+2=4 would be represented in CLIP as **("+" 2 2 | 4)**. In a graphic EG, the last line of identity of the function may be marked by an arrowhead that points from the name **"+"** to the value **4**. The semantics of Common Logic is consistent with Peirce's version of model theory (endoporeutic) for the Alpha and Beta graphs. Whether it's identical to Peirce's intentions for Gamma graphs is debatable. But that issue is independent of using CLIP as a convenient linearization of EGs. In any debate about semantics, Common Logic can serve as a fixed reference point among the many options.

For an overview of the ways Peirce anticipated and often surpassed innovations by his successors, see "Peirce's contributions to the 21st century" [29]. For issues of relating metalanguage and diagrammatic reasoning to modal logic, see "Worlds, models, and descriptions" [30]. The icons in two-dimensional diagrams may be generalized to three dimensions and even 3+1 dimensions for motion and change. The technology of virtual reality can implement them. With VR goggles, people could wander through 3D EGs and manipulate the lines and shapes according to Peirce's rules. Reasoning in any notation requires analogies: the formal method of *unification* is a special case of analogy. To support analogies that map VR imagery to diagrams, the system of Cognitive Memory can find exact and approximate graph matches in logarithmic time [16, 17]. These methods of artificial intelligence, combined with Peirce's logic and semiotic, can serve as a foundation for a cognitive architecture that is compatible with the theories he developed in his late manuscripts. For robots, it can support a kind of *artificial phenomenology*: diagrammatic reasoning with Peirce's rules can map virtual reality to and from language and logic.

References

[1] Arens, Hans (1984) *Aristotle's Theory of Language and its Tradition*, Amsterdam: John Benjamins Publishing Co.

[2] Damasio, Antonio R. (2010) *Self Comes to Mind: Constructing the Conscious Brain*, New York: Pantheon Books.

[3] Deely, John (2013) *Tractatus de Signis: The Semiotic of John Poinsot*, St. Augustine's Press.

[4] Dehaene, Stanislas (2014) *Consciousness and the Brain*, New York: Viking.

[5] Euclid, *The Thirteen Books of the Elements*, translated by Thomas L. Heath, Second Edition, New York: Dover.

[6] Fodor, Jerry A. (1975) *The Language of Thought*, Cambridge, MA: Harvard University Press.

[7] Frege, Gottlob (1879) *Begriffsschrift*, English translation in J. van Heijenoort, ed. (1967) *From Frege to Gödel*, Cambridge, MA: Harvard University Press, pp. 1-82.

[8] Gentzen, Gerhard (1934, 35) Untersuchungen über das logische Schließen, translated by M. E. Szabo as Investigations into logical deduction, *Collected Papers of Gerhard Gentzen*, North-Holland Publishing Co., Amsterdam, 1969.

[9] Hadamard, Jacques (1945) *The Psychology of Invention in the Mathematical Field*, Princeton University Press, Princeton.

[10] Halmos, Paul R. (1968) Mathematics as a creative art, *American Scientist* **56**, 375-389.

[11] ISO/IEC (2007). *Common Logic (CL) — A Framework for a family of Logic-Based Languages*, (IS 24707). Geneva: International Organisation for Standardisation.

[12] Johnson-Laird, Philip N. (2002) Peirce, logic diagrams, and the elementary operations of reasoning, *Thinking and Reasoning* **8:2**, 69-95.

[13] Kamp, Hans, and Uwe Reyle (1993) *From Discourse to Logic*, Dordrecht: Kluwer.

[14] Lakoff, George (1970) Linguistics and natural logic, *Synthese* **22**, 151-271.

[15] Lewis, Clarence Irving (1929) *Mind and the World Order: Outline of a theory of knowledge*, New York: Charles Scribner's Sons.

[16] Majumdar, Arun K., and John F. Sowa (2009) Two paradigms are better than one and multiple paradigms are even better, in S. Rudolph, F. Dau, and S.O. Kuznetsov, eds., *Proceedings of ICCS'09*, LNAI 5662, Springer, pp. 32-47. http://jfsowa.com/pubs/paradigm.pdf

[17] Majumdar, Arun K., and John F. Sowa (2018) Relating language, logic, and imagery, *Postproceedings of BICA 2018*, http://jfsowa.com/pubs/relating.pdf

[18] Peirce, Charles Sanders (1885). On the algebra of logic. *American Journal of Mathematics* 7:180-202.

[19] Peirce, Charles Sanders (1898). *Reasoning and the Logic of Things*, The Cambridge Conferences Lectures of 1898, K. L. Ketner (ed) (1992). Cambridge, MA: Harvard University Press.

[20] Peirce, Charles Sanders (1903) Outline classification of the sciences (CP 1.176 to 1.677), https://www.textlog.de/4258.html

[21] Peirce, Charles Sanders (circa 1906) Phaneroscopy, manuscript R 602, http://jfsowa.com/peirce/ms602.txt

[22] Peirce, Charles Sanders (1931-1958 CP). *Collected Papers of C. S. Peirce* C. Hartshorne, P. Weiss, and A. Burks (eds.), 8 vols. Cambridge, MA: Harvard University Press.

[23] Peirce, Charles Sanders (1976 NEM) *The New Elements of Mathematics*, ed. by Carolyn Eisele, 4 vols., The Hague: Mouton.

[24] Peirce, Charles Sanders (1991-1998 EP) *The Essential Peirce*, ed. by N. Houser, C. Kloesel, and members of the Peirce Edition Project, 2 vols., Indiana University Press, Bloomington.

[25] Pietarinen, Ahti-Veikko (2006) *Signs of Logic: Peircean Themes on the Philosophy of Language, Games, and Communication*, Synthese Library, vol. 329, Berlin: Springer.

[26] Pólya, George (1954) *Mathematics and Plausible Reasoning*, Volume I: *Induction and Analogy in Mathematics*, Volume II: *Patterns of Plausible Inference*, Princeton: University Press.

[27] Putnam, Hilary (1982) Peirce the Logician, *Historia Mathematica* **9**, 290-301.

[28] Roberts, Don D. (1973). *The Existential Graphs of Charles S. Peirce*. The Hague: Mouton.

[29] Sowa, John F. (2006a) Peirce's contributions to the 21st Century, in H. Schärfe, P. Hitzler, and P. Øhrstrøm, eds., *Conceptual Structures: Inspiration and Application*, LNAI 4068, Berlin: Springer, pp. 54-69. http://jfsowa.com/pubs/csp21st.pdf

[30] Sowa, John F. (2006b) Worlds, models, and descriptions, *Studia Logica*, Special Issue *Ways of Worlds II*, **84:2**, 323-360. http://jfsowa.com/pubs/worlds.pdf

[31] Sowa, John F. (2015) Peirce, Polya, and Euclid: Integrating logic, heuristics, and geometry, Lecture slides for a session on Peirce, American Philosophical Association conference in Vancouver, Canada. http://jfsowa.com/talks/ppe.pdf

[32] Sowa, John F. (2017) An introduction to existential graphs. Lecture slides. http://jfsowa.com/talks/egintro.pdf

[33] Sowa, John F. (2018) Reasoning with diagrams and images, *Journal of Applied Logics* **5:5**, 987-1059. http://www.collegepublications.co.uk/downloads/ifcolog00025.pdf

[34] Stjernfelt, Frederik (2007) *Diagrammatology: An Investigation on the Borderlines of Phenomenology, Ontology, and Semiotics*, Berlin: Springer.

[35] Stjernfelt, Frederik (2014) *Natural Propositions: The Actuality of Peirce's Doctrine of Dicisigns*, Boston: Docent Press.

[36] van Benthem, Johan (2008) A brief history of natural logic, In M Chakraborty, B Löwe, M. Nath Mitra and S. Sarukkai (eds.), *Logic, Navya-Nyaya and Applications: Homage to Bimal Krishna Matilal*, London: College Publications.